**Assignment**

To close out this checkpoint, answer the following questions. Once you're done, you can compare your answers to the ones found in [this notebook, which you can download](https://tf-assets-prod.s3.amazonaws.com/tf-curric/data-science/machine-learning-regression/solutions/1.solution_what_is_regression.ipynb).

1. Let's assume that you have World Bank data on financial, economic and social indicators for several countries. You want to measure the factors that affect the level of development in these countries. To this end, you decide to use per capita income as a proxy for the development level, which is defined as the national income divided by the population. You want to use some features in your dataset to predict per capita income. Is this task a classification or a regression task? Why?
2. Which of the following arguments are false and why?
   * OLS is a special type of linear regression models
   * Regression models become useless if they don’t meet the assumptions of linear regression.
   * Estimation and prediction are the same thing in the context of linear regression models.
   * Linear regression is only one kind of regression model. Regression problems can also be solved with other kind of models like Support Vector Machines or Random Forests.
3. Assume that your project manager wants you to discover which free services your company offers make your customers buy more of your paid services. Formulate this task as a regression problem and write down the potential outcome and features that you’d like to work on.
4. This task is regression task. Because the target is “per capita income = income / population”, which is a continuous value. Recall that in contrast to classification problem, “regression problems have a continuous outcome variable”. So, this is a regression task.
   * OLS is a special type of linear regression models. False, “OLS” stands for “Ordinary Least Squares”, which is an optimization algorithm to estimate the coefficients of regression model, it is not a type of linear regression models itself.
   * Regression models become useless if they don’t meet the assumptions of linear regression. False. Since there are “some approaches you can take to use linear regression even when your model doesn’t meet these conditions.”
   * Estimation and prediction are the same thing in the context of linear regression models. False. “estimation” is short for coefficient estimation, as the name indicated, is the process to figure out the coefficients in the linear regression model. While prediction is the process to use the coefficients from the estimation process to predict the outcome given the feature data.
   * Ture
5. Problem formulation: to find out the relationship between paid services sale for each customer and the free services we offered to each customer, so that given the information of free services we offer to each customer, we can predict the service each customer would pay. Outcome: service paid by each customer. Features: free service name, type, the time customer use the free service (duration, frequency), etc.